UNIVERSIDAD CARLOS III DE MADRID working papers

UC3M Working Papers
Statistics and Econometrics
16-01
ISSN 2387-0303
January 2016

Departamento de Estadística
Universidad Carlos III de Madrid
Calle Madrid, 126
28903 Getafe (Spain)
Fax (34) 91 624-98-48

ABC and Hamiltonian Monte-Carlo methods in COGARCH models

Marin J.M.[a], Rodriguez-Bernal M.T.[b], and Romero E.[c]

Abstract
The analysis of financial series, assuming calendar effects and unequally spaced times over continuous time, can be studied by means of COGARCH models based on Lévy processes.
In order to estimate the COGARCH model parameters, we propose to use two different Bayesian approaches. First, we suggest to use a Hamiltonian Montecarlo (HMC) algorithm that improves the performance of standard MCMC methods. Secondly, we introduce an Approximate Bayesian Computational (ABC) methodology which allows to work with analytically infeasible or computationally expensive likelihoods.
After a simulation and comparison study for both methods, HMC and ABC, we apply them to model the behaviour of some NASDAQ time series and we discuss the results.

Keywords: Approximate Bayesian Computation methods (ABC), Bayesian inference, COGARCH model, Continuous-time GARCH process, Hamiltonian Monte Carlo methods (HMC), Lévy process.

[a] Department of Statistics, Universidad Carlos III de Madrid.

[b]Department of Statistics, Universidad Complutense de Madrid.

[c] Department of Economy and Finance, Universidad Europea de Madrid

ABC and Hamiltonian Monte-Carlo methods in COGARCH models

Marín J.M.[1], Rodríguez-Bernal[2] M.T., and Romero E.[3]

(1) Dep. de Estadística, U. Carlos III de Madrid, Spain

(2) Dep. de Estadística e Investigación Operativa, U. Complutense de Madrid, Spain.

(3) Dep. de Economía y Finanzas, U. Europea de Madrid, Spain

Abstract

The analysis of financial series, assuming calendar effects and unequally spaced times over continuous time, can be studied by means of COGARCH models based on Lévy processes.

In order to estimate the COGARCH model parameters, we propose to use two different Bayesian approaches. First, we suggest to use a Hamiltonian Montecarlo (HMC) algorithm that improves the performance of standard MCMC methods. Secondly, we introduce an Approximate Bayesian Computational (ABC) methodology which allows to work with analytically infeasible or computationally expensive likelihoods.

After a simulation and comparison study for both methods, HMC and ABC, we apply them to model the behaviour of some NASDAQ time series and we discuss the results.

Keywords: Approximate Bayesian Computation methods (ABC), Bayesian infer-

ence, COGARCH model, Continuous-time GARCH process, Hamiltonian Monte Carlo methods (HMC), Lévy process.

AMS Subject Classification: C11; C15; C22.

1 Introduction

Financial series present some particular characteristics (see Mandelbrot (1963)) because of the presence of heteroscedasticity, or calendar effects. In such cases, ARCH or GARCH models may be suitable options, but when financial data are observed over unequally spaced times, standard discrete time models may be not adequate.

Although there are different approaches in order to generalize GARCH models on continuous time (see for a review e.g. Kluppelberg et al. (2010) and Granzer (2013)), we will follow in this paper the Klüppelberg et al. (2004) approach. It is based on a single Lévy process which incorporates continuous calendar effects and models the jumps of the process.

Following the Klüppelberg et al. (2004) approach, Maller et al. (2008) characterized a discrete time model which allows to include irregularly spaced observations. They used an embedded sequence of discrete GARCH models $G_n = (G_n(t))_{t\geq 0}$. They also proved that this approximation converges in probability in strong sense to a COGARCH model, with respect to the Skorokhod metric.

Regarding inferential and estimation techniques, Haug et al. (2007) developed a method of moments and they showed that the resulting estimators are consistent and asymptotically normal. On the other hand, Maller et al. (2008) presented estimators based on pseudo-maximum likelihood techniques, and Müller (2010) used a Markov chain Monte Carlo (MCMC) procedure for a COGARCH model driven by a compound Poisson process. Recently, Marín et al. (2015) applied a data cloning

methodology to obtain approximate maximum likelihood estimates by means of a MCMC algorithm.

In regard of a Bayesian point of view, the standard MCMC techniques based on the Gibbs and Metropolis-Hasting samplers may be inefficient in some situations, due to their random walk behaviour. The Hamiltonian Monte Carlo approach (HMC) addresses this situation by defining auxiliary variables and more efficient jumping rules (see Duane et al. (1987)).

On the other hand, also under the Bayesian framework, a new class of algorithms, called Approximate Bayesian Computational (ABC) methods, have been proposed to tackle with likelihoods which are analytically infeasible or computationally expensive. The first type of ABC algorithm was introduced by Pritchard et al. (1999) regarding a Population Genetics problem. The basic idea (see e.g. Csilléry et al. (2010) is based on a rejection algorithm which simulates a large number of datasets under a given parametric model. The parameters are sampled from a probability distribution and reduced to summary statistics; such that the sampled parameters are accepted or rejected if the distance between simulated and observed summary statistics is small enough.

Therefore, for a given a tolerance level and a set of summary statistics, it is possible to obtain a reasonable approximation to the posterior distribution of the model parameters.

In this paper, we deal with a Bayesian approach for inference in COGARCH models based on a Approximate Bayesian Computational (ABC) methodology and we compare its behaviour with the Hybrid Monte Carlo (HMC) technique. In section 2 we revise the main definitions and characteristics of COGARCH models; in section 3 we describe the Hamiltonian Monte Carlo (HMC) and ABC methodologies to obtain Bayesian estimates of the model parameters. In section 4 we first consider

a simulation study of a COGARCH(1,1) model with fixed parameters in order to compare the HMC and the ABC approaches. As results seem to be reasonable in both cases, we apply these methodologies to a real example of the NASDAQ stock index.

2 Introduction to COGARCH Models

In this section, we introduce the definition of the COGARCH model from Klüppelberg et al. (2004) and the approximation in irregular discrete time points proposed by Maller et al. (2008).

Definition 1. Let $(\Omega, \mathcal{F}, P, (\mathcal{F}_t)_{t\geq 0})$ be a filtered probability space, where $(\mathcal{F}_t)_{t\geq 0}$ is the natural filtration of the background driving Lévy process $L = (L_t)_{t\geq 0}$. Given an initial value σ_0, the COGARCH process $G = (G_t)_{t\geq 0}$ and the variance process $\sigma^2 = (\sigma_t^2)_{t\geq 0}$ are defined by the stochastic differential equations

$$dG_t = \sigma_t dL_t \tag{1}$$

$$d\sigma_{t+}^2 = \beta dt - \eta\sigma_t^2 dt + \varphi\sigma_t^2 d[L,L]_t \tag{2}$$

where $t > 0$, $G_0 = 0$, $\beta > 0$, $\eta > 0$, $\varphi \geq 0$ and $[L,L]_t$ is the quadratic variation of the Lévy process. □

Maller et al. (2008) proved that this model can be expressed as a continuous time limit of a sequence of GARCH models. As the COGARCH model may be approximated by an appropriate set of GARCH processes, the parameters of the COGARCH model can be estimated using the relation between them.

In order to deal with the approximation to a COGARCH model, Maller et al. (2008) fixed an interval $[0, T]$, where $T > 0$, and they took the sequence of integers

$(N_n)_{n\geq 1}$ such that $\lim_{n\to\infty} N_n = \infty$, and a sequence of partitions

$$0 = t_0(n) < \cdots < t_{N_n}(n) = T,$$

which divides the interval $[0, T]$ into N_n sub-intervals of length $\Delta t_i(n) = t_i(n) - t_{i-1}(n)$. It is assumed that $\Delta t(n) = \max_{i=1,\ldots,N_n} \Delta t_i(n) \longrightarrow 0$, when $n \to \infty$.

The corresponding discrete-time processes $(G_{i,n})_{i=1,\ldots,N_n}$ and $(\sigma^2_{i,n})_{i=1,\ldots,N_n}$ are defined by

$$\begin{aligned} G_{i,n} &= G_{i-1,n} + \sigma_{i-1,n}\sqrt{\Delta t_i(n)}\epsilon_{i,n} \\ \sigma^2_{i,n} &= \beta\Delta t_i(n) + (1 + \varphi\Delta t_i(n)\epsilon^2_{i,n})e^{-\eta\Delta t_i(n)}\sigma^2_{i-1,n}, \end{aligned}$$

where $i = 1, \ldots, N_n$, and innovations $\epsilon_{i,n}$ are independent and identically distributed with $E(\epsilon_{i,n}) = 0$, $Var(\epsilon_{i,n}) = 1$ and $\sigma^2_{0,n} = \sigma^2_0$.

The continuous-time versions of $G_{i,n}$ and $\sigma^2_{i,n}$ are defined as

$$\begin{aligned} &G_n(t) = G_{i,n} \quad \text{in} \quad t \in [t_{i-1}(n), t_i(n)] \text{ with } G_n(0) = 0 \\ &\sigma_n(t)^2 = \sigma^2_{i,n} \quad \text{in} \quad t \in [t_{i-1}(n), t_i(n)] \end{aligned}$$

If $E(L_1) = 0$ and $E(L_1^2) = 1$, there exists for each $n \geq 1$ a sequence $\epsilon_{i,n}$ such that

$$\rho((G_n, \sigma^2_n), (G, \sigma^2)) \xrightarrow[n\to\infty]{P} 0,$$

where ρ is the *Skorokhod* distance.

Therefore, the process $(G_n(t), \sigma_n(t))$ can be viewed as an approximation to the COGARCH model $(G(t), \sigma(t))$ for n large enough.

Maller et al. (2008) used this approximation to fit the model to unequally spaced time data, and they derived a pseudo-maximum likelihood function and obtained

the estimates of the corresponding parameters by numerical maximization. We will address the previous pseudo-likelihood function as the basis of the joint posterior distribution of the parameters in the HMC approach.

Accordingly, let us assume that data $G(t_i)$ are distributed as a COGARCH model, defined in (1) and (2), in time points $0 = t_0 < t_1 < \cdots < t_N = T$. Let $Y_i = G(t_i) - G(t_{i-1})$ be returns and $\Delta t_i = t_i - t_{i-1}$ the time increments. As the process $(\sigma_{t_i})_{i=1\ldots,N}$ is Markovian, Y_i is conditionally independent of the previous returns Y_{i-1}, $Y_{i-2}, \ldots$ given $\mathcal{F}_{t_{i-1}}$.

The conditional expectation and variance of Y_i given $\mathcal{F}_{t_{i-1}}$, by using Proposition 5.1 in Klüppelberg et al. (2004), are

$$\begin{aligned} E(Y_i|\mathcal{F}_{t_{i-1}}) &= 0 \\ \rho_i^2 &= Var(Y_i|\mathcal{F}_{t_{i-1}}) = E(Y_i^2|\mathcal{F}_{t_{i-1}}) = \\ &\quad \left(\sigma_{t_{i-1}}^2 - \frac{\beta}{\eta-\varphi}\right)\left(\frac{e^{(\eta-\varphi)\Delta t_i}-1}{\eta-\varphi}\right) + \frac{\beta\Delta t_i}{\eta-\varphi}, \end{aligned} \tag{3}$$

where $(L_t)_{t\geq 0}$ is a quadratic pure jump process with $E(L_1) = 0$, $E(L_1^2) < \infty$ and $(\sigma_t^2)_{t\geq 0}$ is stationary.

To ensure the stationarity of the process it must be taken on that $E(L_1^2) = 1$ and $E(\sigma_0^2) = \frac{\beta}{\eta-\varphi}$, with $\eta > \varphi$.

Then, assuming that Y_i are conditionally $N(0, \rho_i^2)$, the pseudo-likelihood function for the observed returns $\mathbf{y} = (y_1, \ldots, y_N)$ is

$$L(\beta, \varphi, \eta|\mathbf{y}) = \left(\prod_{i=1}^{N} \frac{1}{\sqrt{2\pi\rho_i^2}}\right) \exp\left\{-\sum_{i=1}^{N} \frac{y_i^2}{2\rho_i^2}\right\},$$

where

$$\begin{aligned}
\sigma_i^2 &= \beta\Delta t_i + e^{-\eta\Delta t_i}\sigma_{i-1}^2 + \varphi e^{-\eta\Delta t_i}(\sqrt{\Delta t_i}\epsilon_{i,n}\sigma_{i-1})^2 \\
&= \beta\Delta t_i + e^{-\eta\Delta t_i}\sigma_{i-1}^2 + \varphi e^{-\eta\Delta t_i} y_i^2 \qquad (4) \\
\sigma_0^2 &= \frac{\beta}{\eta - \varphi}
\end{aligned}$$

Equation (4) is obtained by substituting $\sigma_{t_{i-1}}^2$ by σ_{i-1}^2 in (3), and assuming the discretization of the volatility process.

3 ABC and HMC estimation in COGARCH models

Bayesian methods have been recently applied in COGARCH models by means of a MCMC approach (see e.g. Müller (2010)) and a data cloning approach (see Marín et al. (2015)).

In this section, we also follow a Bayesian approach in COGARCH models; in particular, we address a Hamiltonian Monte Carlo (HMC) and an approximate Bayesian computational (ABC) methodologies. HMC methodology improves the performance of a standard MCMC approach, and ABC is a simulation-based approach that allows to work with cumbersome likelihoods functions. In this section, we study both techniques and we compare their behaviours with respect to COGARCH(1,1) model.

3.1 Hamiltonian Monte Carlo estimation in COGARCH models

In a Bayesian framework, computation times can be large, particularly in cases like COGARCH models given the complexity of these processes. In this sense, a Hamiltonian Monte Carlo approach provides more efficient Metropolis-Hasting steps which allow to obtain a more accurate and faster approximation to the posterior distributions of parameters.

The original idea of HMC methods is based on Hamiltonian dynamics (see e.g. Gelman et al. (2014) and Neal (2011) for a survey) as a general technique in theoretical Physics, where the total energy of a system of particles is described by means of a function called *Hamiltonian function*.

When these ideas are applied in the context of Bayesian Inference, the position of particles can be seen as the parameters of interest, i.e. $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_d)$, and the potential energy of particles can be understood as their log–posterior distributions.

An auxiliary *momentum* variable ϕ_j $(j = 1, \ldots, d)$ is added for each component θ_j, such that both members of the couple are jointly updated for each j. Accordingly, the posterior distribution of $\boldsymbol{\theta}$ is *augmented* by the distribution of the auxiliary parameter $\boldsymbol{\phi}$, in order to obtain the joint posterior distribution $P(\boldsymbol{\theta}, \boldsymbol{\phi}|y)$.

In the HMC procedure, simulations are derived from $P(\boldsymbol{\theta}, \boldsymbol{\phi}|y)$, although $\boldsymbol{\phi}$ is just an auxiliary variable that enables the algorithm to move faster in the parameter space. Additionally, in the procedure, it is required to calculate the gradient of the log–posterior distributions which, in practice, it can be computed by means of numerical differentiation.

Now, we first describe the HMC approach for a COGARCH(1,1) model assuming the approximation proposed by Maller et al. (2008) and described in section 2. The

model parameters of interest are β, η and φ which play the role of *position* variables in the HMC scheme.

We set as prior distributions, vaguely-informative uniform distributions,

$$\beta \sim U(0, a), \quad \varphi \sim U(0, c), \quad \eta|\varphi \sim U(\varphi, b) \tag{5}$$

where a, c and b are positive real values and $b > c$.

Then, the posterior distribution is

$$\begin{aligned} & \pi(\beta, \varphi, \eta|y_1, \cdots, y_N) \propto L(\beta, \varphi, \eta)\pi(\beta)\pi(\varphi)\pi(\eta|\varphi) \\ \propto & \left(\prod_{i=1}^{N} \frac{1}{\sqrt{2\pi\rho_i^2}}\right) \exp\left\{-\sum_{i=1}^{N} \frac{y_i^2}{2\rho_i^2}\right\} \cdot I_{(0,a)}(\beta) \cdot I_{(0,c)}(\varphi) \cdot I_{(\varphi,b)}(\eta|\varphi) \\ \propto & \left(\prod_{i=1}^{N} \frac{1}{\rho_i}\right) \exp\left\{-\sum_{i=1}^{N} \frac{y_i^2}{2\rho_i^2}\right\} \cdot I_{(0,a)}(\beta) \cdot I_{(0,c)}(\varphi) \cdot I_{(\varphi,b)}(\eta|\varphi) \end{aligned} \tag{6}$$

and ρ_i is defined as

$$\rho_i^2 = \left(\sigma_{i-1}^2 - \frac{\beta}{\eta - \varphi}\right)\left(\frac{e^{(\eta-\varphi)\Delta t_i} - 1}{\eta - \varphi}\right) + \frac{\beta \Delta t_i}{\eta - \varphi},$$

where $\sigma_i^2 = \beta\Delta t_i + e^{-\eta\Delta t_i}\sigma_{i-1}^2 + \varphi e^{-\eta\Delta t_i} y_i^2$ and $\sigma_0^2 = \frac{\beta}{\eta-\varphi}$, with $\eta > \varphi$.

The potential energy is the derivative of the log-posterior density function,

$$\frac{d \log \pi\left(\beta, \eta, \varphi|y_1, \cdots, y_N\right)}{d\left(\beta, \eta, \varphi\right)}.$$

The momentum variables (ϕ_1, ϕ_2, ϕ_3) are auxiliary variables that are assumed to be distributed as a normal distribution $\phi_j \sim N\left(0, 1\right)$ for $j = 1, 2, 3$.

3.2 ABC estimation of COGARCH models

Approximate Bayesian Computational (ABC) methods can be deemed as a natural way to find estimates in models, where the likelihood function is not completely known or it is intractable, although it is possible to simulate from it.

In ABC algorithms, samples are simulated from a given model and they are compared with observed data. A more operative approach (instead of a direct comparison between the real data and the simulated samples), is to use a statistic ξ and to fix a small enough distance between them, bounded by a term ϵ.

Pritchard et al. (1999) proposed the first practical ABC algorithm: Assume we observe some data $\mathbf{y} \in \mathcal{D} \subset \mathbb{R}^n$, then we define:

i) A set of summary statistics $\xi(Y)$ (a function on data but usually not a sufficient statistic).

ii) A distance ρ between the the simulated and the real data by means of statistic ξ.

iii) A tolerance level ϵ.

Then, the algorithm can be summarized as follows:

for $i = 1$ **to** N **do**

– **Repeat**

— Generate θ' from the prior distribution $\pi(\cdot)$

— Generate $\mathbf{z}$ from the likelihood $f(\cdot|\theta')$

– **until** $\rho\{\xi(\mathbf{z}), \xi(\mathbf{y})\} \leq \varepsilon$

Set $\theta_i = \theta'$

end for

The algorithm of Pritchard et al. (1999) obtains samples of the joint distribution

of $\boldsymbol{\theta}$ and $\mathbf{z}$, from the marginal distribution of $\mathbf{z}$,

$$\pi_{\varepsilon}(\boldsymbol{\theta}, \mathbf{z}|\mathbf{y}) = \frac{\pi(\boldsymbol{\theta}) f(\mathbf{z}|\boldsymbol{\theta}) I_{A_{\epsilon,y}}(\mathbf{z})}{\int_{A_{\epsilon,y}\times\theta} \pi(\boldsymbol{\theta}) f(\mathbf{z}|\boldsymbol{\theta}) d\mathbf{z} d\boldsymbol{\theta}},$$

where $I(\cdot)$ is an indicator function and

$$A_{\epsilon,y} = \{\mathbf{z} \in D | \rho\{\xi(\mathbf{z}), \xi(\mathbf{y})\} \leq \epsilon\}.$$

Then, the posterior distribution is approximated as

$$\pi_{\epsilon}(\boldsymbol{\theta}|\mathbf{y}) = \int \pi_{\epsilon}(\boldsymbol{\theta}, \mathbf{z}|\mathbf{y}) d\mathbf{z} \approx \pi(\boldsymbol{\theta}|\mathbf{y}).$$

Noticeably, when the tolerance level ε becomes smaller better approximations to the posterior distribution are obtained, but at a higher computational cost.

The accurate behaviour ABC algorithm depends on the appropriate selection of the tuning parameters: the summary statistics ξ, the distance ρ and the tolerance ϵ. These terms have to be set before running the algorithm, and handy selections of them improve the results in this methodology. Although an universal procedure for the calibration and setting of the tuning parameters remains up today as an open issue.

McKinley et al. (2009) carried out an empirical study to test different strategies to select the tolerance level ϵ, the distance ρ and the summary statistics ξ. They noticed that the selection of summary statistics and distance is essential for obtaining accurate approximations of posterior distributions.

In order to implement an ABC algorithm to estimate the COGARCH model parameters, we have tested some possible sets of summary statistics ξ and distances ρ.

With regards of ρ, the standard euclidean distance seems to work appropriately. Apropos of the set of statistics ξ, we include several relevant issues of financial series:

i) The volatility lower bound, because parameters η and β are related with it.

ii) The magnitude of the process jumps (difference between the maximum and the minimum observed volatilities), that is related with parameter φ.

iii) The correlation coefficient of the squared values, related with the autocorrelation of the squared returns of financial series.

iv) The sample variance.

v) The median absolute deviation.

Finally, we have noticed that, in order to find good enough results in a reasonable computational time, it is better to limit the prior distributions of parameters around the pseudo–maximum likelihood, or the moments estimates of them. Accordingly, we have used as prior distributions of parameters, uniform distributions whose supports are located around the moment estimates (see Haug et al. (2005)).

In the next Section, the performances of the HMC and ABC approaches are explored under simulated time series. Then, both procedures are applied in a real data example.

4 Applications of HMC and ABC methods

In this section, we first undertake a simulation study to check and to compare the estimates obtained by the HMC and ABC methodologies in a COGARCH(1,1) model (see section 3).

Then, we apply both methods in order to model some Nasdaq index returns (from January 2008 to December 2012) and the corresponding results are compared.

4.1 HMC and ABC estimations for COGARCH(1,1) simulated data

In this simulation study, we have simulated a series of 100 data points with parameters $\beta = 0.25$, $\eta = 0.35$ and $\varphi = 0.02$.

In the HMC method we follow the approach of section 3.1. Here, we assume uniform prior distributions as in (5), where $a = 100, b = 10$ and $c = 0.05$. Then, we approximate the posterior distributions of parameters (6) by simulating 3 chains with 5000 iterations, each one.

The HMC algorithm have been programmed in `Stan` (Stan Development Team (2014b)) software by means of `Rstan` package (Stan Development Team (2014a)).

In the ABC methodology, we follow the procedure shown in section 3.2. The algorithm is programmed in `R` (R Core Team (2012)) and codes were run with parallel computing by means of library `doParallel` (see Analytics and Weston (2014)) in order to diminish computing times in multicore machines.

We use uniform prior distributions (5) but we bound their domains to close intervals near the moments estimates of the parameters (see Haug et al. (2005)), by using the package `COGARCH` (Bibbona et al. (2014)) in order to calculate them. Then, we use the package `ABC` (Csillery et al. (2012)) to analyse the corresponding outputs. We have run 100000 simulations with a tolerance level of 0.01, in such a way that 1000 sets of parameters have been accepted as a sample from the posterior distributions.

The results obtained by applying the HMC and ABC algorithms in the simulated data set are shown in table 1. It displays the real values of parameters, the obtained estimates, the standard deviations and the HPD intervals.

Regarding the accuracy of HMC and ABC methods, it may be noticed that for

Estimates using ABC				
Parameter	Real Value	Posterior mean	S.D.	HPD 0.95
β	0.25	0.335	0.119	(0.118, 0.501)
η	0.35	0.297	0.145	(0.098, 0.569)
φ	0.02	0.024	0.014	(0.003, 0.047)
Estimates using HMC				
Parameter	Real Value	Posterior mean	S.D.	HPD 0.95
β	0.25	0.542	0.202	(0.162, 0.948)
η	0.35	0.323	0.144	(0.088, 0.367)
φ	0.02	0.027	0.014	(0.002, 0.048)

Table 1: Estimation of COGARCH(1,1) parameters using ABC and HMC methods

all parameters estimates are close to their corresponding real values (remarkably η and φ), and they are all included in the HPD intervals. Furthermore, for parameter β the proposed ABC algorithm obtains a smaller HPD interval.

Although both methods obtain quite similar results, the practical advantage of the ABC approach is that it is a pure simulation-based method, where it is not necessary to deal with approximate quasi–maximum likelihood estimation or methods of moments. Although, the computational burden may be huge if the support of the prior distributions is not restricted somehow, as we have considered in section 3.2.

4.2 HMC and ABC estimations for a NASDAQ-100 stock index data set

In this section, we apply a COGARCH(1,1) model to analyse the behaviour of the Nasdaq-100 stock index from January 2000 to November 2012. Data have been obtained from the website `finance.yahoo.com`. In order to estimate the parameters, we apply both HMC and ABC methodologies.

As in Section 4.1, in the HMC methodology we assume uniform prior distributions (5), where $a = 100, b = 10$ and $c = 0.05$. Then, we approximate the posterior

distributions of parameters (6) by simulating 3 chains with 5000 iterations each one. Correspondingly, the HMC algorithm have been programmed in `Stan` (Stan Development Team (2014b)) software by means of `Rstan` package (Stan Development Team (2014a)).

In the ABC procedure, we have used 100000 simulations with a tolerance level of 0.01, then a sample of 1000 sets of parameters have been accepted from the posterior distribution. We have undertaken parallel computing with codes by means of library `doParallel` (see Analytics and Weston (2014)) to optimize computing times.

As in Section 4.1, we use uniform prior distributions (5), whose domains are bounded close to the moments estimates of parameters (see Haug et al. (2005)). We use the package `COGARCH` (Bibbona et al. (2014)) in order to calculate the moments estimators, and we use the package `ABC` (Csillery et al. (2012)) to analyse the outputs of the program.

Results obtained with HMC can be compared with those obtained by an ABC approach. Table 2 shows similar posterior means, standard errors and HPD intervals for the COGARCH(1,1) parameters. Although with a long time series as in this case, the computing times of ABC are much larger than in the HMC approach, rounding ten times more time (3 days with a 4 cores i7 computer).

Estimates using ABC			
Parameter	Posterior means	S.D.	HPD 0.95
β	9.663	0.379	$(9.040, 10.198)$
η	0.077	0.031	$(0.051, 0.136)$
φ	0.033	0.013	$(0.007, 0.049)$
Estimates using HMC			
Parameter	Posterior means	S.D.	HPD 0.95
β	9.637	2.064	$(6.155, 14.289)$
η	0.053	0.004	$(0.045, 0.062)$
φ	0.046	0.002	$(0.038, 0.049)$

Table 2: ABC and HMC estimates for a COGARCH(1,1) model for Nasdaq daily returns data

In general, parameter η measures the speed of the decline of a volatility burst, and for this dataset the value is not very high, namely, when a volatility burst appears due to the arrival of new information to markets, its influence in volatility declines with a moderate speed.

By the other hand, parameter φ measures the magnitude of the volatility jumps and it may be viewed as a measure of how information affects to volatility and how fast that the market assumes new events. In this Nasdaq-100 series the estimate of φ is not large, so volatility bursts seem to be small and daily index are stable regarding the appearance of relevant information in market.

Finally, parameter β represents the level of volatility and regarding this Nasdaq-100 series, the estimated value is quite large.

5 Final conclusions

Nowadays, GARCH modelling is a very popular methodology, that takes into account the most important stylized facts that financial series present, and it is applied in equally spaced data. But there are periods of time with holidays and weekends that may affect the behaviour of series, and it is convenient to introduce unequally spaced or continuous time periods. In this sense, COGARCH models may be a good alternative to tackle with this kind of situations.

In this work, we use COGARCH models based in Lévy processes which include volatility and model the process jumps (see Klüppelberg et al. (2004)). We propose to use Hamiltonian Monte Carlo (HMC) and Approximate Bayesian Computation (ABC) methodologies to estimate their parameters.

The main obtained results are:

i) Both HMC and ABC approaches can be used for inference in COGARCH

models as they show a good behaviour in simulation studies.

ii) We have analysed the behaviour of the Nasdaq-100 index series under a COGARCH(1,1) model, by means of HMC and ABC methodologies. Results are quite similar in both cases and they can be easily interpreted from a practical point of view.

iii) Although the ABC method may be used in long real time series, and programs can be efficiently parallelized, the computing times are larger than in the case of the HMC approach. So in this point HMC would be a better option.

References

Analytics, R. and S. Weston (2014). *doParallel: Foreach parallel adaptor for the parallel package*. R package version 1.0.8.

Bibbona, E., I. Negri, A. Arcagni, and L. Mercuri (2014). *COGARCH: Simulation and inferenc for COGARCH(1,1) processes*. R package version 1.3-1/r32.

Csilléry, K., M. G. Blum, O. E. Gaggiotti, and O. François (2010). Approximate bayesian computation (ABC) in practice. *Trends in ecology & evolution 25*(7), 410–418.

Csillery, K., O. Francois, and M. G. B. Blum (2012). abc: an R package for approximate bayesian computation (ABC). *Methods in Ecology and Evolution*.

Duane, S., A. D. Kennedy, B. J. Pendleton, and D. Roweth (1987). Hybrid Monte Carlo. *Physics letters B 195*(2), 216–222.

Gelman, A., J. B. Carlin, H. S. Stern, and D. B. Rubin (2014). *Bayesian data analysis*, Volume 2. Taylor & Francis.

Granzer, M. (2013). Estimation of COGARCH models with implementation in R. Master's thesis, Technische Universität München.

Haug, S., C. Klüppelberg, A. Lindner, and M. Zapp (2005). Estimating the COGARCH(1,1) model-a first go.

Haug, S., C. Klüppelberg, A. Lindner, and M. Zapp (2007). Method of moment estimation in the COGARCH(1,1) model. *The Econometrics Journal 10*(2), 320–341.

Klüppelberg, C., A. Lindner, and R. Maller (2004). A continuous-time GARCH process driven by a Lévy process: Stationarity and second-order behaviour. *Journal of Applied Probability 41*(3), 601–622.

Kluppelberg, C., R. Maller, and A. Szimayer (2010). The COGARCH: a review, with news on option pricing and statistical inference.

Maller, R. A., G. Müller, A. Szimayer, et al. (2008). GARCH modelling in continuous time for irregularly spaced time series data. *Bernoulli 14*(2), 519–542.

Mandelbrot, B. (1963). The variation of certain speculative prices. *The journal of business 36*(4), 394–419.

Marín, J. M., M. T. Rodríguez-Bernal, and E. Romero (2015). Data cloning estimation of GARCH and COGARCH models. *Journal of Statistical Computation and Simulation 85*(9), 1818–1831.

McKinley, T., A. R. Cook, and R. Deardon (2009). Inference in epidemic models without likelihoods. *The International Journal of Biostatistics 5*(1).

Müller, G. (2010). MCMC estimation of the COGARCH(1,1) model. *Journal of Financial Econometrics 8*(4), 481–510.

Neal, R. (2011). MCMC using Hamiltonian dynamics. *Handbook of Markov Chain Monte Carlo*, 113–162.

Pritchard, J. K., M. T. Seielstad, A. Perez-Lezaun, and M. W. Feldman (1999). Population growth of human y chromosomes: a study of y chromosome microsatellites. *Molecular Biology and Evolution 16*(12), 1791–1798.

R Core Team (2012). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. ISBN 3-900051-07-0.

Stan Development Team (2014a). Rstan: the R interface to Stan, version 2.5.0.

Stan Development Team (2014b). *Stan Modeling Language Users Guide and Reference Manual, Version 2.5.0*.